Molecular basis of plant viral virulence; the complete nucleotide sequence of an attenuated strain of tobacco mosaic virus

Masamichi Nishiguchi*, Shoshi Kikuchi, Yukio Kiho, Takeshi Ohno[1], Tetsuo Meshi[2] and Yoshimi Okada[2]

National Institute of Agrobiological Resources, Tsukuba Science City Yatabe, Ibaraki 305, [1]Department of Agricultural Chemistry, Faculty of Agriculture, Tokyo University of Agriculture and Technology, Fuchu-shi, Tokyo 183, and [2]Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, Tokyo 113, Japan

ABSTRACT

The total genome sequence of $L_{11}A$, an attenuated strain of tobacco mosaic virus (TMV), has been determined. This strain is able to multiply in tomato plants without inducing any remarkable symptoms, but to protect them from later infection with virulent TMV strains. When compared with the recently published total genome sequence of TMV L (the virulent ancestral strain of $L_{11}A$) ten base substitutions were found in the $L_{11}A$ genome. Seven of these occurred in the third letters of in-phase codons and did not influence amino acids. Only three, which were in the common reading frame for both the 130K and 180K proteins, resulted in amino acid changes. Together with the result of the partial sequence of RNA of L_{11}, an intermediate strain in sequential isolation from L to $L_{11}A$, it is observed that one base at the nucleotide position 1117 is changed from L to L_{11} and two bases at the positions 2349 and 2754 are changed from L_{11} to $L_{11}A$.

INTRODUCTION

An attenuated strain of tobacco mosaic virus (TMV), $L_{11}A$, has been used in Japan to protect tomato plants against wild virulent strains (1). L A hardly produces any symptoms on them (2). This strain, then, would seem to be suitable material for the study of the virulence of TMV. Some of its properties have been reported (2,3,4), however no information suggestive of its molecular mechanism of attenuation has been obtained. In this study we determined the complete nucleotide sequence of $L_{11}A$ RNA and compared it with that of L, the ancestral virulent strain (5).

MATERIALS AND METHODS

Materials---The TMV tomato strain ($L_{11}A$) was inoculated on Nicotiana tabacum cv. Samsun after six series of single local lesion isolation using N.tabacum cv. Xanthi nc and Lycopersicon esculentum cv. Fukuju no.2. Virus was purified(6) and its RNA was extracted as described previously (7,8). Ω-RNA fragments were prepared as described (9). Reverse transcriptase was purchased from Life Science, Inc. Restriction enzymes were obtained from

Takara Shuzo Co., New England Biolabs, BRL, and Nippon Gene Co. dNTP's and ddNTP's were from Boehringer Mannheim Yamanouchi, Inc.

Cloning and sequencing---The cDNA copies of TMV-L_{11}A genomic RNA were cloned by the method of Okayama and Berg (10) after in vitro polyadenylation of L_{11}A RNA as described before (11). The cDNA clones were screened by colony hybridization. Genomic RNA fragments sometimes in combination with Ω-RNA fragments, were used as hybridization probes (11). The sequences of cDNA clones were determined by the method of Maxam and Gilbert (12) and partly by the method of Sanger et al.(13). RNA was directly sequenced using the dideoxy sequencing method modified for reverse transcription of an RNA template as described before (14).

RESULTS AND DISCUSSION

The whole set of the cloned cDNAs were found to cover the entire genome of L_{11}A and used for sequence analysis. Because it could not be determined that L_{11}A genomic RNA is clonally pure even though it has been passed through many series of single local region isolation, at least two different clones were sequenced for a given region. Fig. 1 shows the sequencing strategy of each clone. The restriction patterns of each clone indicate that pL_{11}A-A25 possibly carried cDNA covering the whole genomic RNA and that pL_{11}A-2-35 lacked only about 230 5' terminal nucleotides. cDNAs in these two clones were partly sequenced. Shorter cDNAs cloned independently in pL_{11}A-A97, pL_{11}A-2-10 and pL_{11}A-2-16 were fully sequenced. The genome of L_{11}A was composed of 6383 nucleotides. At least four open frames were deduced in the L_{11}A genome as well as in both the L (5) and vulgare genomes (15). Compared with the nucleotide sequence of the L genome (see Fig. 3 of ref.5), the L_{11}A genome was one base shorter. This deletion occurred in the A cluster (6336-6341), which consisted of 6 adenine residues in the 3' non-coding region of L_{11}A. Though the published number of adenine residues is 7, a heterogeneity ranging from 6 residues to 9 was observed in the A cluster of the L genome when genomic RNA was directly sequenced (11). There was no heterogeneity in that region of the L_{11}A genome. This might be attributable to a serial single local region transfer before purifying L_{11}A virus. As shown in Fig. 2, ten base substitutions were found to be dispersed over the genome. All of the base substitutions were of the transition type. The amino acid sequences of the 30K and coat proteins of L_{11}A were the same as those of L, since all four base substitutions occurring in these coding frames were in the third letters

Fig. 1 The cloned cDNAs of L_{11}A genomic RNA and the strategy used for the sequence analysis. The sequence of cDNA was determined by the method of Maxam and Gilbert (12) (●——➤) and partly that of Sanger et al. (13) (○——➤). Terminal dots of arrows indicate the labeled end of DNA fragments for the method of Maxam and Gilbert or the primer positions for that of Sanger et al. The sites for restriction enzymes are shown only to indicating the positions of terminal dots of arrows.

Fig. 2 Distribution of mutations of the L_{11}A genome and the proteins coded in it. The position of the nucleotide is shown by the nucleotide number from the 5' terminus of the genome. (ⵀ) indicates the nucleotide change that does not affect the amino acid change and (ⵀ) indicates the change that affects the amino acid change. Nucleotides and amino acids are shown in parentheses and in brackets, respectively. Lower nucleotides are for L and upper for L_{11}A. At the positions of 1117, 2349 and 2754, the nucleotides and the amino acids in L_{11} genome are also shown in the middle.

in-phase codons. The assembly origin of L is located within the 30K protein cistron (16,17). The origin is thought to exist as a presumed hairpin loop structure formed by bases 5440-5530 (11). One base difference between L and $L_{11}A$ occurred in this region (at 5475) but not specially in the target sequence (-GAPuGUUG-; 5479-5485) (18,19). Thus, this substitution probably would not have any effect on the assembly of the virion. In the cistrons of 180K and 130K, however, three base substitutions from G of L to A of $L_{11}A$ at 1117, from A to G at 2349 and from G to A at 2754 resulted in the changes of amino acids Cys, Asn and Gly to Tyr, Asp and Arg of $L_{11}A$ respectively.

In order to get further information about the relationship between these three base substitutions and changes in virulence, we compared the sequence of L_{11} (2,20) at the positions of 1117, 2349 and 2754 with those of L and $L_{11}A$. L_{11} is an attenuated strain isolated by Oshima et al.(20) from L-inoculated tomato stems incubated at 35°C for two weeks using the method of Holmes (21). L_{11} induces very mild symptoms in tomato plants in the later stage of infection whereas $L_{11}A$ does not even at that stage. A part of the genomic RNA of L_{11} was sequenced using the modified dideoxy method described in Materials and Methods. The base at 1117 was A in both the L_{11} and $L_{11}A$ genomes (Fig. 3), whereas the bases at 2349 and 2754 were A and G respectively, in both L_{11} and L (data not shown). Since the
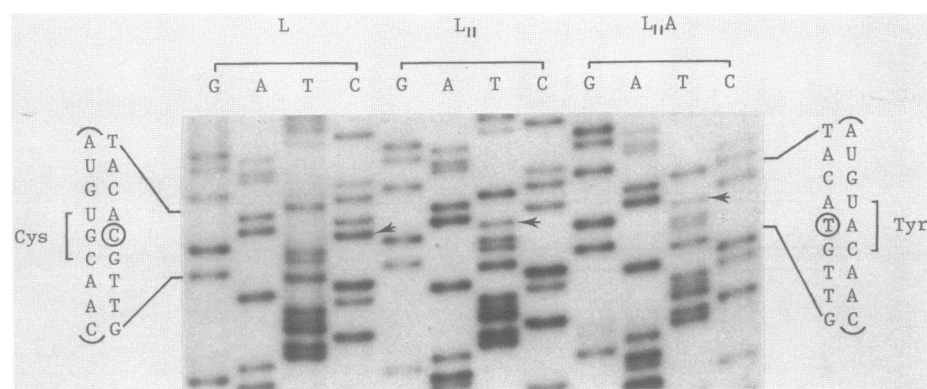

Fig. 3 Autoradiogram of the RNA-sequencing gel of the region with a single base substitution at position 1117 that leads to an amino acid change. The reverse transcripts were separated by electrophoresis through a 6% polyacrylamide gel containing 53% (w/v) urea. Arrowheads indicate the nucleotides where a substitution occurred. Note that the complementary nucleotide sequences are shown in this modified method.

virulence of L_{11} is much closer to that of $L_{11}A$ than that of L, it seems that the base substitution at 1117 is more contributive to the attenuation of virulence in the process of the sequential selection of $L_{11}A$ via L from L. The deduced amino acid sequence around the nucleotide position 1117 (amino acid position 348 of the 130K and 180K proteins) is well conserved when this region was compared between L and vulgare (5). This suggests that this region is crucial for the functions of the 130K and 180K proteins. On the other hand, amino acid sequences corresponding to the region at 2349 and 2754 (amino acid positions 759 and 894) differ between L and vulgare, implying that this region is flexible and strain specific. Considering that disulphide bonds between cysteine residues play important roles in the functioning of some proteins (22), it seems that the change of cysteine of L to tyrosine of $L_{11}A$ at amino acid position 348 causes a modification of the tertiary structure of 130/180K proteins and results in the change in biological activity. The amino acid change in these proteins appears important in attenuation of virulence, but in this case it is difficult to determine that only the amino acid change at the position 348 affects the viral virulence, because the possibility cannot be excluded that other amino acids would be concerned. To elucidate this problem other attenuated strains should be analysed. It is not clear what the function of 130K and 180K proteins is. There is some evidence to suggest that these proteins are involved in the replication of viral RNA (23,24). Considering this point of view it is interesting that the virus yield of L A in the host plant is about one-fifth less than that of L (25). This low productivity may be due to an autoregulation mechanism of $L_{11}A$ (26) which was presumed to exist from the observation that the rate of multiplication of $L_{11}A$ decreased drastically in the later stage of infection, while L still continued to multiply steadily. It remains to be elucidated what the mechanism of autoregulation is and how the altered 130K and 180K proteins participate in this mechanism.

Our results contrast with a report concerning an attenuated strain of polio virus, whose critical difference from virulent strain exists in the coding frames of virus structural proteins (27). $L_{11}A$ differed from L in the coding frame of non-structural protein.

REFERENCES
1. Oshima, N. (1981) Japan Agric. Res. Quart. 14, 222-228.
2. Goto, T. and Nemoto, M. (1971) Hokkaido Natl. Agric. Exp. Stn. Bull. 99, 67-76
3. Nishiguchi, M. and Oshima, N. (1979) Ann. Phytopath. Soc. Japan 45, 654-659.
4. Nagai, Y. (1981) Spec. Bull. Chiba Agric. Exp. Stn. 9, 1-109.
5. Ohno, T., Aoyagi, M., Yamanashi, Y., Saito,H., Ikawa, S., Meshi, T. and Okada, Y. (1984) J. Biochem. 96, 1915-1923.
6. Otsuki, Y., Takebe, I., Ohno, T., Fukuda, M. and Okada, Y. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 1813-1817.
7. Gierer, A. and Schramm, G. (1956) Nature 177, 702-703.
8. Frankel-Conrat, H., Singer, B. and Tsugita, A. (1961) Virology 14, 54-58.
9. Mandels, S. (1968) J. Biol. Chem. 243, 3671-3674.
10. Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
11. Takamatsu, N., Ohno, T., Meshi, T. and Okada, Y. (1983) Nucleic Acids Res. 11, 3767-3778.
12. Maxam, A. M. and Gilbert, W. (1980) Methods in Enzymol. 65, 499-559.
13. Sanger, F.M., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
14. Meshi, T., Ishikawa, M., Takamatsu, N., Ohno, T. and Okada Y. (1983) FEBS Lett. 162, 282-285.
15. Goelet, P., Lomonossoff, G. P., Butler, D.J.G., Akam, M.E., Gait, M. J. and Karn, J. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 5818-5822.
16. Zimmern, D. and Hunter T. (1983) The EMBO J. 2, 1893-1900.
17. Zimmern, D. (1983) The EMBO J. 2, 1901-1907.
18. Meshi, T., Ohno, T., Iba, H. and Okada, Y. (1981). Mol. Gen. Genet. 184, 20-25.
19. Meshi, T., Kiyama, R., Ohno, T. and Okada, Y. (1983). Virology 127, 54-64.
20. Oshima, N., Komochi, S. and Goto, T. (1965) Hokkaido Natl. Agric. Exp. Stn. Bull. 85, 23-33.
21. Holmes, F. O. (1934) Phytopathology 24, 845-873.
22. Mainwarring, W. I. P., Parish, J. H., Pickering, J. D. and Mann, N. H. (1982) in Nucleic Acid Biochemistry and Molecular Biology, 434-435. (Blackwell, Oxford)
23. Dawson, W. O. and White, J. L. (1979) Virology 93, 104-110.
24. Sakai, F. and Takebe, I. (1974) Virology 62, 426-433.
25. Nishiguchi, M., Motoyoshi, F. and Oshima, N. (1981) Ann. Phytopath. Soc. Japan 47, 421.
26. Kiho. Y. and Nishiguchi, M. (1984) Microbiol. Immuno. 28, 589-599
27. Nomoto, A., Omata. T., Toyoda, H., Kuge, S., Horie, H., Kataoka, Y., Genba, Y., Nakano, Y. and Imura, N. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 5793-5797.